#2½

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Malcolm J. Simons

Assignee: GeneType AG

Title: "INTRON SEQUENCE ANALYSIS METHOD FOR DETECTION OF ADJACENT AND REMOTE LOCUS ALLELES AS HAPLOTYPES"

Serial No. ~~unknown~~ 07/949652 Filed: herewith

Examiner: ~~unknown~~ PAUL B. FLUM Group Art Unit: 1807

Attorney Docket No.: M-1647-6C

THE COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, DC 20231

DECLARATION

Sir:

I, Peter Gresshoff, hereby declare the following. I am Professor, Racheff Chair of Excellence, Plant Molecular Genetics, University of Tennessee. Attached is a copy of my curriculum vitae.

I previously submitted a Declaration regarding Malcolm Simons' discovery that one could use relatively short regions of non-coding sequences, on the order of one to two kilobasepairs, to define the corresponding coding region allele. At that time, Malcolm Simons had definitively demonstrated that relatively short non-coding region sequences contained informative polymorphisms which can be used as the basis of a complete HLA typing system. That demonstration was not something I would have expected from the literature on the subject or my own research, and I was impressed by the technology.

At that time, I felt that the phenomenon probably related to genes generally. However, because the phenomenon had been demonstrated in the HLA genes, a family of genes which are

LAW OFFICES OF
SKJERVEN, MORRILL,
MacPHERSON, FRANKLIN
& FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233

prone to biological diversity or to genes which have a high degree of coding region polymorphism, there was a question in my mind as to whether this phenomenon could be related to members of the immunoglobulin super-gene family or to gene families with high coding region variability.

Since that time, I have become convinced that the basis for Malcolm Simons' analysis system applies to all eukaryotic genomes. That is, I have seen additional data for other phylogenetically distant genes which convinced me that the phenomenon that non-coding region sequences contain informative polymorphisms that can be used to identify associated coding region alleles is a general phenomenon.

More specifically, as part of my research, I study the nitrogen tolerant symbiosis (NTS) gene (also referred to as the supernodulation gene) which is present in several species of plants, notably soybeans. This gene regulates a plant's ability to fix nitrogen from the atmosphere in a symbiotic relationship with bacteria in nodules on the roots of the plant, making the plant independent of nitrogen in the soil. The most common wild-type alleles of the gene in soybeans, for example as found in the Bragg cultivar, have a feedback mechanism which turns off nodule formation when the soil is rich in nitrogen. Another wild-type allele of the gene, found in the Williams cultivar, also turns off nitrogen fixation even when the soil is rich in nitrogen. We have isolated mutant alleles, which confer both supernodulation and nitrate-tolerant symbiosis. This lack of inhibition of nitrogen fixation by the presence of nitrogen in the soil, referred to as super-nodulation, is commercially beneficial because one can use those soybeans to enrich the soil prior to rotating in a crop which is not capable of fixing nitrogen such as corn or wheat.

LAW OFFICES OF
SKJERVEN, MORRILL,
MacPHERSON, FRANKLIN
& FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233

Since the trait is commercially beneficial, breeders try to obtain varieties of soybeans which have alleles of the NTS gene which do not turn off nitrogen fixation. However, there is currently no simple genetic test to determine which allele of the NTS gene has been inherited. The exact location of the supernodulation gene is unknown. As part of my research, I have been analyzing portions of genomic DNA in regions near the NTS gene to find a way to distinguish co-cultivars of soybeans to develop a typing system for the NTS gene. I identified a region of DNA sufficiently near the gene so that recombination is very unlikely. From the types of nucleotides present in the region and the lack of an open reading frame, it is my opinion that this region is an intergenic region (70% AT). That is, the region is located between genetic loci, rather than within a genetic locus.

Using standard methods (including an RFLP-type analysis of amplified sequences of this region of DNA in which the fragments from a restriction endonuclease digestion of the amplified sequence were analyzed by gel electrophoresis to determine differences in the length of the fragments), I was unable to find any differences between the co-cultivars. Nevertheless, I sequenced portions of these apparently invariant DNA regions which I had amplified and analyzed by standard methods. In sequencing DNA from the soybean co-cultivars which I had previously analyzed, I found that there were VNTRs, site-specific polymorphisms, and minimum deletions in a region as long as one kilobase. That is, I found heterogeneity on the sequence level (micro-heterogeneity) where restriction endonuclease digestion revealed no differences in the sequences. I also determined that some of these non-coding region polymorphisms were indicative of the

LAW OFFICES OF
SKJERVEN, MORRILL,
MacPHERSON, FRANKLIN
& FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233

co-cultivar. That is, the non-coding region polymorphism could be used to determine whether the DNA in the sample was from the Bragg or the Williams cultivar. These findings convinced me that relatively short non-coding region sequences contained informative polymorphisms which can be used as the basis of a typing system for genes generally.

I found these additional data convincing for a number of reasons. The first is that because we simply were working in this non-coding, presumably intergenic region of the soybean genome and found this pattern of variation, I feel certain that these same patterns of variation must be present in other parts of the soybean genome and in the genomes of eukaryotes, generally. That is, we did not perform these studies to determine whether Malcolm Simons' work applied to plants. In addition, we have no reason to believe this region is anything other than typical of intergenic regions of the soybean genome or of the genomes of other plants. Therefore, since this region which appeared invariant by standard analyses exhibited micro-heterogeneity, there is no reason to expect that other regions of the soybean genome or any other plant genome would be different.

In addition, although we have not sequenced the NTS gene, we believe it is a conserved gene from its mutation rate. Yet the same correlation of non-coding region polymorphisms with coding region polymorphisms which is present in the HLA genes is also present in the soybean NTS gene. In addition to demonstrations of this non-coding region micro-heterogeneity in both highly polymorphic and presumably conserved genes, the data were obtained in humans and in soybeans. Clearly, this indicates that the phenomenon is not limited to humans or even animals.

LAW OFFICES OF
SKJERVEN, MORRILL,
MacPHERSON, FRANKLIN
& FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233

What appears to be happening is that random changes which occur in the sequence of the coding region (which give rise to new phenotypes and alleles) are correlated with changes in the sequence of the non-coding region. The probability of undoing that correlated change in the non-coding region is so small that the correlation remains over evolutionary time periods. Or it is feasible that intergenic regions (as well as intragenic non-translated regions like introns) possess properties which influence either the expression of neighboring genes (enhancers, silences, endo-duplication origins) or the stability and adaptability of the genome. As such, they would be suject to selection and co-inheritence association. When one looks at the allelic divergence between the human races which took place over 30,000 generations, the non-coding region changes associated with the coding region changes are still present as indicated by Malcolm Simons' HLA data. This same type of association of changes in the non-coding region which correlate with coding region changes is also present in soybeans, so that we can distinguish alleles of the NTS gene even from very closely related soybean co-cultivars.

Since the analysis method is applicable to both the HLA loci in humans and the supernodulation locus in soybean, in my opinion, the approach is applicable to eukaryotes generally.

LAW OFFICES OF
SKJERVEN, MORRILL,
MacPHERSON, FRANKLIN
& FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: Sept 2, 1997

Peter Gresshoff

I hereby certify that this correspondence is being deposited with the United States Postal Service as express mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C., 20231, on _____ 19__. Express Mail Receipt No. RB 605 849 706 US

9/23/97
Date of Signature

LAW OFFICES OF
SKJERVEN, MORRILL, MACPHERSON, FRANKLIN & FRIEL
25 METRO DRIVE
SUITE 700
SAN JOSE, CA 95110
(408) 283-1222
FAX (408) 283-1233